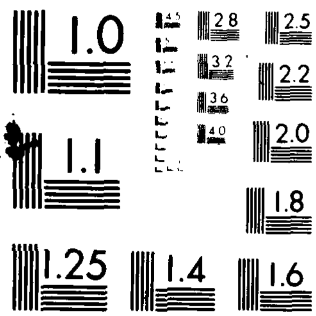
AD-A110 247 L N K CORP SILVER SPRING MD F/G 9/2
ANALYSIS AND SIMULATION OF DISCRETE DIGITAL IMAGE MATCHING.(U)
NOV 81 D LAVINE, B A LAMBIRD, L N KANAL DACA76-81-C-0004
UNCLASSIFIED ETL-0278 NL

END
DATE
FILMED
2 82
DTIC



MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS-1963-A

LEVEL II

ETL-0278

12

Analysis and Simulation of discrete digital image matching

David Lavine

Barbara A. Lambird

Laveen N. Kanal

L.N.K. Corporation
302 Notley Court
Silver Spring, MD 20904

NOVEMBER 1981

U.S. ARMY CORPS OF ENGINEERS
ENGINEER TOPOGRAPHIC LABORATORIES
FORT BELVOIR, VIRGINIA 22060

ETL

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER ETL-0278	2. GOVT ACCESSION NO. AD-A110 247	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Analysis and Simulation of Discrete Digital Image Matching		5. TYPE OF REPORT & PERIOD COVERED Contract Report 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) David Lavine Barbara A. Lambird Laveen N. Kanal		8. CONTRACT OR GRANT NUMBER(s) DACA76-81-C-0004
9. PERFORMING ORGANIZATION NAME AND ADDRESS L.N.K. Corporation 302 Notley Court Silver Spring, MD 20904		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS Program Element 6.27.07-A Project No. 4A762707A855 Task Area B Work Unit No. 00026
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Engineer Topographic Laboratories Fort Belvoir, VA 22060		12. REPORT DATE November 1981 13. NUMBER OF PAGES 82
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

digital image registration, feature extraction, correlation, image matching, simulation, point features, clustering.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

This report evaluates procedures for the approximate registration of digital images to maps and other images. It emphasizes point features and vectors derived from point features for image registration. Labeling and other means of adding definiteness to feature identification are promoted as ways to speed up registration. Clustering algorithms, distribution and type of features, hierarchical matching techniques, and other procedures are investigated with scene and map simulations.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 391118

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

Analysis and Simulation of Discrete Digital Image Matching

Table of Contents

PREFACE

The work presented in this report was performed by L.N.K. Corporation under Contract DACA76-81-C-0004 for the U.S. Army Engineer Topographic Laboratories, Fort Belvoir, Virginia 20060. The contract technical monitor was Mr. A. T. Blackburn. The authors thank Mr. Blackburn and Mr. Mike Crombie for helpful interactions and technical reviews during the performance of this work.

1. Introduction

The research presented in this report extends earlier work by the L.N.K. Corporation [Stockman 1981], [Lambird 1980], and [Stockman 1979]. The registration of an image to a map or an image to an image, possibly from different sensors, is treated in this report. The LNK registration procedure is designed to handle these tasks in a fast and reliable manner. The present form of the algorithm is the result of a long development process in which both real edges and feature points such as road intersections played a role. Due to the simplicity and reliability of feature points, we have focused on the analysis of their use in registration.

Registration using point features consists of two facets. First, curves and straight lines must be extracted from the image in order to locate high curvature points and intersections. Second, the extracted feature points must be connected to form abstract vectors and the registration must be performed with the abstract vectors.

The present study consisted of further algorithm development for feature point detection and the analysis of the registration procedure, given that a set of feature points has been selected. The algorithm development consisted of the design and implementation of a state-space search procedure for the merging of short curve segments into more reliable curves coupled with an intersection detector for curves.

The analysis of the registration procedure treated a variety of aspects of the problems including abstract vector selection procedures, abstract triples of points rather than vectors, vector labelling procedures, and registration accuracy. These results are based on both analytic and simulation models. While these results do not comprise a complete picture of algorithm performance, they have proven useful in designing modifications to the existing algorithm.

The ultimate usefulness of the analysis will not be clearly established until the intersection procedures have been applied

to a substantial number of pictures in order to obtain a variety of point images. The theory and simulations provide insight into the manner in which such a study should be conducted. In particular, the intersection detection procedure requires further testing to determine the probability of prominent features being detected.

Experience with the LNK procedure has led us to the conclusion that for a wide range of numbers of image feature points, the primary problem in attaining an accurate registration is to guarantee an adequate number of correctly corresponding abstract vectors without incurring a heavy computational cost. Thus, unless a large number of noise points are present, non-corresponding points do not tend to reduce registration accuracy. In light of these observations, much of our work has been directed towards procedures for obtaining an adequate number of matching vectors or triples, rather than the local accuracy of the procedure. In the correlation literature, for these two aspects of registration, the probability of being roughly correct and the probability of being very accurate given that you are roughly correct are treated independently.

The results of the registration analysis indicate that a combination of sampling of abstract vectors, using triples of feature points and labelling feature points as to type, such as a straight road intersection in a field, can reduce the number of feasible registration transformations sufficiently to allow rapid checking of the remaining possibilities. If this conclusion is corroborated by experimentation on real data, the clustering portion of the algorithm can be discarded, thus eliminating a costly and somewhat unpredictable portion of the program.

The cost of feature extraction is high, making it desirable to process only a portion of the image if possible. The analysis of sampling procedures described in the report can be extended to provide design criteria for a sequential registration program which examines only as much of the image as is necessary to determine the registration transformation.

2. Basic L.N.K. Registration Technique

This section presents an overview of the basic LNK registration procedure. The goal of registration is to use automatically extracted features from two images (or an image and a map) to find the best global transformation that maps one image to the other.

The LNK procedure is able to achieve this even when many local mismatches of features occur. The three basic steps of the method are: 1) Primitive features, such as line segments, edge segments, intersections, or high curvature points are automatically extracted. These features may be parameterized, such as by position, by length, or by orientation. 2) Assume all features of one type can correspond to one another. For example, an intersection of three lines in one image can correspond to any three-line intersection in the second image. However, only one of these correspondences is the true one. Now, for each of the possible correspondences, find the translation and rotation that maps one feature to the other. Let that translation and rotation transformation be represented by the triple (θ, x, y). Place a unit of weight in the bin in the three-dimensional histogram that represents (θ, x, y). This process is repeated until all possible feature correspondences have been transformed. 3) Locate any prominent clusters of (θ, x, y) in the histogram. Each cluster represents a set of features in one image that could be mapped to corresponding features in the other image by one particular (θ, x, y). The best global transformation is defined by the (θ, x, y) of the largest cluster. This transformation provides the largest number of local correspondences.

This method works because the correct transformation will show up as a large cluster, while the wrong transformations will tend to be distributed randomly throughout the histogram. The procedure may also be performed iteratively by making the bin sizes smaller and smaller.

As stated in Step 1 of the procedure, many features could be used to perform the transformation. Not all features are equally desirable. For example, the lengths of edge segments or line segments usually cannot be accurately determined. Small perspective changes can alter the curvature of high curvature points considerably. LNK feels that intersections would produce more accurate transformations as they can be determined more accurately. LNK has performed enough experiments to show that even if up to 90 percent of the detected features in the image do not match with any features in the map, the correct transformation will still form enough of a cluster to be detected.

All of the versions of LNK's registration technique require clustering to be done in θ-space. The cluster space is filled with points, each one of which represents the matching of one image element to one map element on the basis of local features alone. A cluster of points in this space represents a globally good transformation that matches many image elements to corresponding map elements. A threshold may be set to determine the strength of an acceptable cluster; for instance, we might demand that half of the map elements be matched by image elements. A unique cluster will typically result, but there may be no cluster formed in the case where there is poor feature detection or where the image and map really do not match. On the other hand, several strong clusters can result as in the case where symmetry produces several good matching possibilities. LNK has used two different clustering techniques - hierarchical clustering and variable resolution clustering.

In the hierarchical technique, clustering for $= (\theta, x, y)$ was done, first on θ alone, and then in (x, y)-space given a fixed θ. This technique was useful for human interaction since θ-space could be viewed as a histogram and (x, y)-space as a scatter plot.

The θ-space was divided into 360 one-degree bins and the rotation $\Delta\theta$ that rotated the image edge to the map edge was entered. After smoothing, the top three peaks of the histogram

were chosen for the second step. For each of the transformations that had a $\Delta\theta$ in one of the three peaks, the (x, y) in that transformation was recorded in (x, y)-space. The resulting scatter plot could then be searched for clusters.

Intuitively, clustering can be done by moving a box around the cluster space to see if there is some position at which an acceptable number N of points are inside the box. If so, the coordinates (θ, x, y) of that position mark a cluster center and represent a good registration transformation. A computer implementation can be specified when we choose the size of the box and the set of different positions (bins) at which we will try to fill it.

There is one final note on the clustering procedure. Since the cluster space is quantized by the bins, overlapping sets of bins are actually used so that clusters are not lost by being split by bin boundaries. Thus each triple (θ, x, y) is actually distributed to several bins rather than one.

To use detected straight edges instead of abstract edges, some changes need to be made in the clustering algorithm. Real edge detectors create edges with two general defects. First, only small segments of the true edge are usually detected primarily because only a fixed size detector is used. Second, the detectors may overshoot the true termination of an edge at a corner.

In order to deal with the above problems when real edges are used, the matching of image vectors to map vectors must be made to be sloppy. The image edge element must be allowed to slide along the potentially matching model edge. Mathematically this generates an infinite set of triples (θ, x, y) constrained to be on a line segment in cluster space. In practice the endpoints (θ, x_1, y_1) and (θ, x_2, y_2) are in separate bins, points are also contributed to all bins in between them.

The work in this report focuses on the situation where abstract vectors rather than real edges are used, so some of the considerations outlined above will not play a direct role in the following analysis. In Sections 4 and 5 we consider various

procedures which may allow us to dispense with the clustering entirely.

5. Curve Intersection Detection

The detection of intersections of straight line segments in imagery may not provide an adequate supply of feature points for the LNK registration procedure. Even if a substantial number of such intersections are found, the straight-line intersections may have few distinguishing characteristics. To supplement or possibly replace these intersections, the intersections of curves may be used. This more general procedure is likely to be useful in images which contain little or no man-made imagery, hence few straight lines. Curved intersections may be labelled by a feature vector describing the curvature of the intersecting curves. This feature vector may then be used to greatly reduce the number of potential matches between image and feature points.

An algorithm to perform this curve detection is given in this section. Testing of this algorithm on aerial imagery will be necessary to evaluate it properly. The formation of curves from local evidence of edge content can be a costly procedure due to the many ways in which short curved segments can be joined to form longer curves. Experimentation on aerial images should be performed to estimate some bounds on the computational requirements for a range of pictures, as well as estimating the potential benefits.

An extension of the curve intersection detector to intersections of paired parallel curves, such as roads, is currently under consideration. This restricted type of intersection detection may substantially reduce the number of rotation and translation transformations which must be considered in the latter part of the LNK procedure.

3.1 Continuous Curve Extraction

The extraction of continuous curves consists of two operations; the detection of edge points and the linking of edge points into continuous curves. The details of this procedure are given in [Stockman 1981]. We review this material briefly to provide insight into the input of the curve intersection program described in Section 3.2.

A point can be discovered to be an edge point by testing the tonal values in halfplanes on either side of the point. Figure 3.1 shows the 32 directions for edges that were used in the reported research. A general purpose routine (RPSML) exists which can differentiate either 8, 16, or 32 directions around the circle. Three "masks" for computing directional gradient values are shown in Figure 3.2. Direction d = 1 is the vertical direction with higher tones at the right while direction d = 7 is nearly horizontal with higher tones below. Given a pixel (x,y) in the image, the gradient magnitude in each direction d = 1,2,. . .,32 can be computed by adding and subtracting tonal values as indicated by the masks in Figure 3.2. Division by a normalization factor is performed to take into consideration the number of pixels used and to make the magnitude geometrically isotropic. The magnitude and direction of the gradient at (x,y) is taken to be the magnitude and direction where a maximum is achieved. Note that for d = 17 the mask for d = 1 could be used with reversal of the sign on the magnitude, so that only 16 masks are actually applied. Only 4 directions are tried at resolution 8. The points selected by the masks sometimes differ from the ideal due to implementation considerations.

The presence of one edge may interfere with the detection of neighboring edges. Suppression of known edge points may be required in this situation.

By examining the neighbors of edge pixels, it is sometimes possible to determne the two pixels which continue the edge in the forward and backward direction. (Forward edge traversal by definition keeps higher tonal values toward the right.)

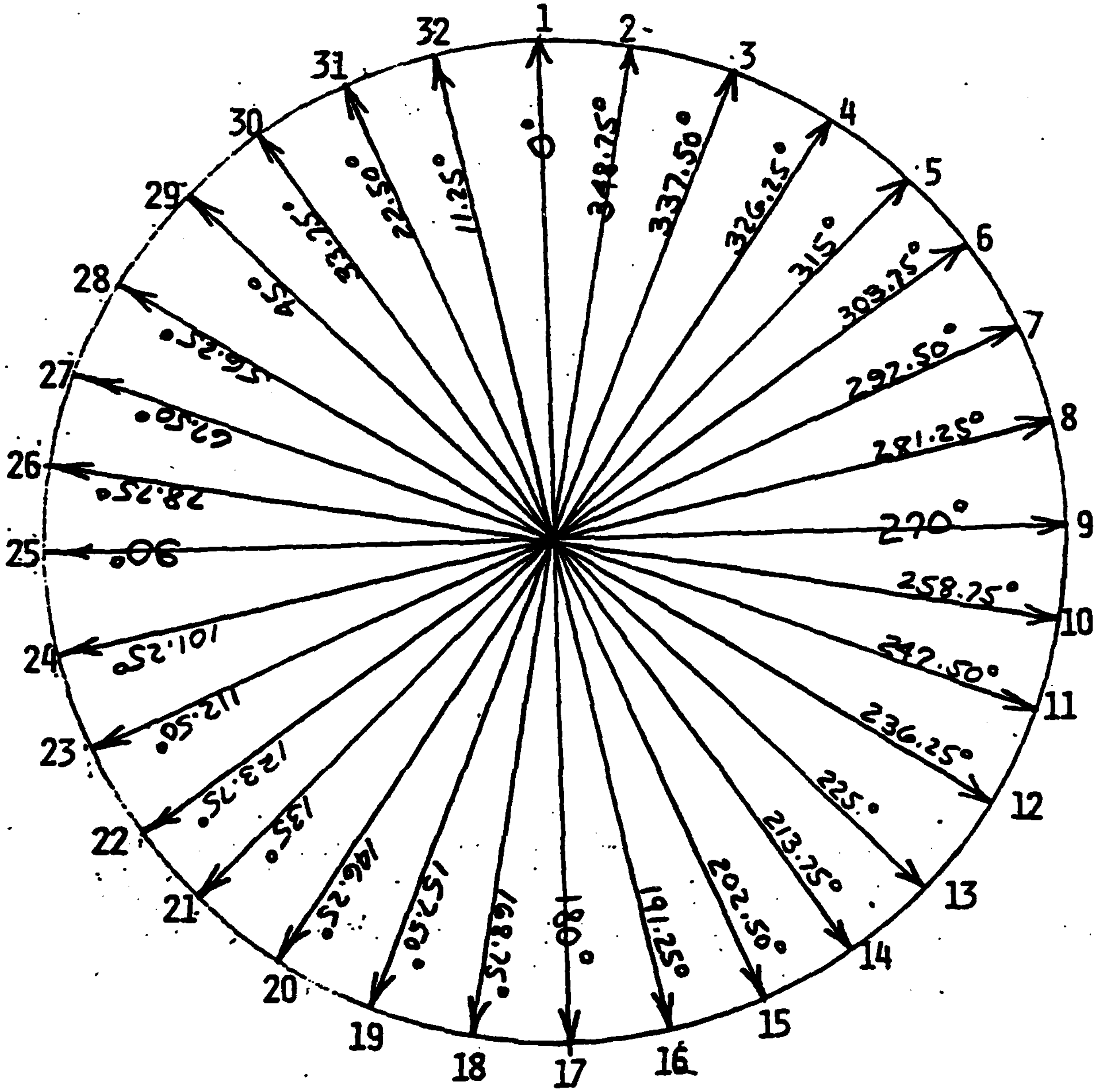


Figure 3.1 32-directional codes and associated angular heading in degrees.

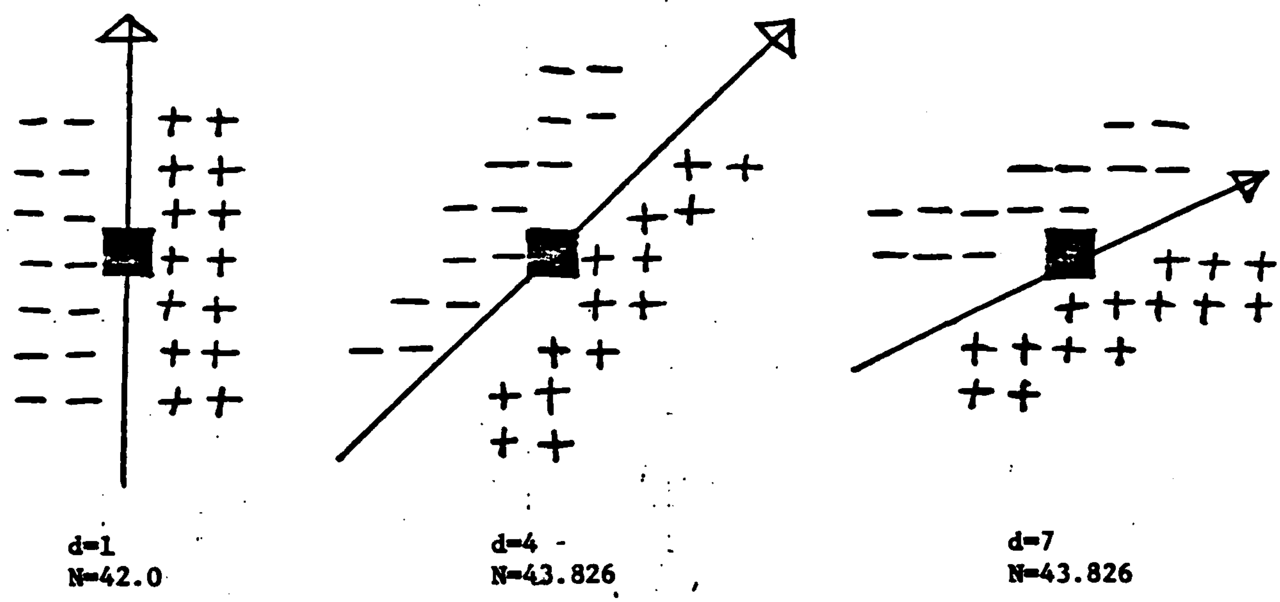


Figure 3.2 Three masks used for computing the gradient at a point in directions 1,4, and 7. The directional resolution is one 32nd of the circle; N is the normalization factor. Edge directions of 1,4, and 7 are indicated by the arrows.

Continuation can be determined from the spatial orientation of the neighbors and the gradient direction of the neighbors all relative to that of the original pixel. When determining which neighbor is the best forward and backward continuation, a variable-sized neighborhood is scanned in a spiral pattern. See [Stockman, 1979] for a detailed discussion of continuous edge linking.

3.2 Curve Intersection Detection Algorithm

This section describes an algorithm for the detection of curves and the subsequent detection of intersections. This algorithm has been implemented and undergone limited testing. Further testing and algorithm refinement are now being considered.

The first phase of curve detection is the computation of the gradient and linking as described in Section 3.1.

The next stage in the intersection detection process is to perform a piecewise linear approximation to the curve. A sequential procedure for extracting the linear approximation is performed. The algorithm sequentially examines the straight-line segments joining one end of the curve to its neighbor, joining the end to the third point on the curve, etc. For each line segment, the sum of squared error between the curve and the segment are computed. As soon as the error exceeds a threshold, the line segment joining the previous point to the endpoint of the curve is selected.

The error associated to each point on the curve is the vertical or horizontal distance from the point to the straight line segment. For purposes of numerical stability, the horizontal distance is chosen if the absolute value of the slope is greater than one and the vertical distance is chosen otherwise. The sums of squared error is the sum of the squared distance for each point on the curve to the corresponding point on the line segment.

Having found one line approximating part of the curve, the algorithm repeats the process using the endpoint of the just-fitted line segment interior to the curve as the new starting endpoint. The procedure repeats until the whole curve is fitted. Though this procedure guarantees no global optimum least squares error fit to the entire curve, it has performed well in practice and is much easier than global optimization procedures.

The next stage in the intersection detection process is the

merging of curves. A hierarchical merging procedure is now invoked for finding the longest possible curves that can be formed from shorter curves which fit well together. This merging procedure employs an optimization procedure known as state-space search to find good merges. Initially, a list of pairs of curves which are candidates for merges is set up. Each pair of curves in the image has associated with it a set of merits indicating the quality of the potential merging. At present, four measures of merge merit are calculated and combined to form a single merit figure. Only curve pairs with a merit figure less than a threshold are placed on the list of merge pairs, called the open list for merges. Any pairs of curves which have an endpoint of the piecewise linear approximation of one curve close to an endpoint of the piecewise linear approximation of the other curve, but which fails to satisfy the merge criterion, is placed on a list. This list, called the intersecton open list, contains curves which are likely candidates for intersections.

Processing proceeds by removing from the merge open list the curve pair with the highest merit. Let A and B denote the curves and let A have endpoints a_1 and a_2 and let B have endpoints b_1 and b_2 and assume a_2 and b_1 are the close endpoints. Then a new symbol, C, denoting the merging of those curves, is created. Any curve D, such that (A,D) is on the open list and a is the endpoint near D results in the pair (C,D) being put on the merge open list. Similarly, any curve E such that (D,E) is on the open list and b_2 is the vertex near E, results in the pair (C,E) being put on the merge open list.

The merging of curves may be intuitively thought of as a soldering of the wires which are close together. The above condition for forming new pairs can be interpreted as saying we first solder two wires together and then allow this soldered wire to form a pair with any wire near its free ends but not with wires near the soldered joint. For each new pair placed on the open list, a set of merge merits is computed.

A pair of curves which are merged may result in a new curve which has no other curves with which it may be merged. This can

happen in two ways. First the new curve may not have either endpoint near the endpoint of any other curve. In this case the curve is removed from the open list and plays no further role in the merge process. Second, the curve may be near an endpoint of another curve, but the new curve pair may not satisfy the merge criterion. If this occurs, the pair is put on the intersection open list.

The process of removing the best pair from the open list and generating merged curves and possibly new pairs for the merge list or the intersection list is repeated until the merge open list is empty. At this point the intersection open list contains the list of curve pairs which represent possible intersections.

The merit function for merging pairs is based on four measures of merit, namely distance, alignment, good continuation, and strength. Each merit measure assumes values between zero and one. The four merit measures are added and the sum divided by four to yield a merit function taking values from zero to one, where lower values indicate better merges.

Distance, the first of the merit measures, is defined to be the distance between the nearby endpoints of the curves divided by the maximum length of the curves. By the distance between nearby endpoints, we refer to the quantity

$$d = \min\ (d(a_1, b_2),\ d(a_2, b_1))$$

where one curve has endpoints (a_1, a_2), the second curve has endpoints (b_1, b_2), and d denotes Euclidean distance. An alternate distance measure may be defined by using the average of the curve lengths rather than the maximum but this has not yet been explored.

Alignment refers to the difference in direction between the two close ends of the curves which may be merged. The line segments on the close ends of the curves are used for this measure. Those line segments have directions induced by the direction of the underlying curves. Directions from zero to 360 are quantized in 32 levels and assigned a value as in Figure 3.1.

The continuation merit of a curve merge is defined to be the average of the angles formed by the end line segments of the

curves with the line joining their nearby end points. If the curves align properly, each angle is zero. Finally, the strength of a curve is a measure of the intensity of the image on the curve. For a line segment in the piecewise linear approximation to a curve, we define the strength of the line segment to be the average grey level for points on the curve between the endpoints of the line segment. The strength of a curve is thus defined to be the average strength merit of its line segments. We then define the strength of a curve pair to be one minus the maximum of the strength of the two curves divided by the maximum grey level in the picture.

4. Registration Analysis

The analysis of the LNK registration procedure involves a study of various sampling questions and in one approximation to the problem, an analysis of discrete image correlation. The results of this section represent our efforts to treat those aspects of the problem we consider most relevant to registration accuracy. We were unable to find results in the literature providing accuracy of similar point matching problems though related point matching techniques have appeared [Fischler and Bolles 1981].

Through much of Section 4, we focus on the study of screening procedures for eliminating abstract vectors which do not contribute to potential registration transformations. This allows processing of a substantial percentage of the set of all possible image vectors, thus insuring that many correct matching vectors will be given to the registration procedure.

4.1 Point Image Modelling

The modelling of point feature images is useful in the understanding of the LNK registration procedure. At this time, a primary practical value of this modelling is in the design of simulation studies for the algorithm, although some analytical results are presented. Though the successes of image modelling have often been limited by the inadequacies of the models in representing real-world problems, we think the simplicity of the point image input to the abstract vector registration procedure permits the development of applicable models.

Before describing any formal image models, we discuss those features of the real-world problem which we think are likely to affect the performance of the registration procedure. The models will deal only with images consisting of a small number of points, typically less than 100, of uniform intensity on a background of a uniform but different intensity. Due to the dependence of the number of points on both image content and the feature extraction procedure, this number is difficult to estimate. We expect, in many cases, the image to contain far fewer than 100 points. The map may contain only a sample of the feature points occurring on a real map, perhaps as few as five or ten. For simplicity, we normalize the points to have intensity one and the background to have intensity zero.

No attempt has yet been made to analyze the probability of feature points occurring given a description of the original grey scale image. Such an analysis would treat questions such as the probability of a road intersection detector failing to detect an intersection if one were present, and of falsely detecting an intersection when none was present. This type of analysis will not be feasible, even on a purely empirical basis, until we have had more opportunity to study feature selection procedures. A second type of empirical study of use in modelling would be the estimation of the distribution of feature points on maps and

real-world imagery. Here we would be concerned with human selection of features that we would like our feature point detectors to find. Though this type of study would involve complex sampling problems, it would aid greatly in refining our registration models.

Our use of the term "point image modelling" has been loose, and we now wish to clarify matters. Two types of matching problems are considered in this report: image-to-image matching and image-to-map matching. The analytic formulation of the two problems is similar, but in the case of image-to-map matching, we shall often be able to make simplifying assumptions due to our confidence that map features are reliable. As a result, we shall use the term "point image" to refer to an image formed from either a real image or a map.

The essential idea we wish to capture in our point image modelling relates to the types of image noise. Consider first the case of matching an image to a map. We assume that certain points have been selected from the map to serve as feature points. The images will contain, in general, some points corresponding to the map points and some additional points. At this point, several difficulties are present in the problem. First, not all image points correspond to map points, and vice versa (by correspond we mean, "correspond to the same ground features"). Second, corresponding image and map points may not have the same positions due to distortions induced by elevation and feature extraction inaccuracies. A generating mechanism for creating such images is to create a set of map points which have corresponding image points. Next create a copy of these points in the image but perturb the location of each one. Finally, generate points randomly in the map and another set of points randomly in the image representing the non-corresponding points.

The procedure described above for generating image-map pairs can be applied to generating image pairs. Some significant

differences are likely to exist in the final models, for by tuning feature extraction procedures, we may be confident that most map points will occur in the image. However, there may be many image points not occurring in the map. Thus the image map problem may be simplified by assuming no noisy points exist in the map. This same assumption also allows us to select a small number of map ponts in comparison with the number of image points (see Section 4.7).

The model, as stated above, requires the specification of probability distributions for generating corresponding points, perturbations of corresponding points, and background noise points. A simple mathematical model arises if we generate corresponding points and background noise points uniformly and perturb corresponding points using bivariate Gaussian noise. Since this model is the basis of many of our simulation studies, the plausibility of these assumptions will now be discussed.

The assumption that corresponding points are generated by a bivariate uniform distribution is, in many cases, clearly inappropriate. Urban areas have many feature points such as road intersections which are reglarly located. For such areas the assumption of uniformly distributed corresponding points may lead to overly optimistic performance predictions. For more rural areas and areas in which natural rather than man-made point features are selected, the assumption of uniformity may be more tenable.

The assumption of uniformity for the distribution of corresponding points may be dropped if background points are uniformly generated and we assume we are matching an image to a map and the map has a sharply peaked autocorrelation function. Under these assumptions the autocorrelation of the map to itself will not contribute significantly to registration error. Thus registration error will result primarily from high peaks in the cross-correlation function of the map with the background noise.

Since the map and the noise are uncorrelated, we expect high peaks to be rare, though this topic requires further investigation.

The above observations have an interesting potential application. By applying the registration procedure to two copies of the map and comparing the match weight at the time of transformation with the match weight at other peaks selected by the registration procedure, we have a method of assessing the extent to which the structure of the map is likely to lead to confusion. If it were then possible to establish the expected matching between background noise points and the map points, then criteria could be set for determining whether or not a successful match has been made. Such criteria would be extremely useful in iterative feature extraction and registration procedures (see Section 6) in which the decision as to whether or not to do further feature extraction is based on the success of the registration.

The assumption that image or noise points are uniformly generated may be replaced by the assumption of a Poisson distribution for the points. Observing that the uniform distribution of a fixed number of points is the same as a poisson distribution conditioned on the number of points, the difference between these two models is not great. We have chosen the former model since the number of image and map points will be known at the time of registration, thus allowing for more accurate performance predictons. With a Poisson model, the performance predicton is averaged over various numbers of points occurring. A further limitation of the Poisson model is that it does not allow us to control the mean and variance independently. Thus, for instance, as we increase the expected number of map points, we increase the variance of the number of map points forcing us to average performance prediction over a wide range of numbers of map points.

The Poisson image model is a limiting case of a binomial

image model in which an image is formed by generating a zero or one at each pixel according to a binomial distribution. This type of model is similar to the Gaussian model used for gray scale images in which each pixel is assumed to have an intensity generated from a Gaussian distribution.

Just as the binomial image model may be viewed as a discrete version of the Poisson process image model, a set of points uniformly distributed in a finite array of pixels may be viewed as the discrete analog of the continuous image model with a fixed number of points uniformly distributed.

All simulations and analyses in this report are based upon the assumption that the number of points is fixed and that the points have coordinates generated from a uniform random number generator. In experimental situations, the number of map points is small, generally ranging from five to ten while the number of image points ranged from zero to fifty, depending upon the experiment. These values were selected for several reasons. First, as stated above, the map used for matching can be sparsely sampled from a true match, so we are free to choose as few points as possible as long as we can get a correct transformation. Second, by placing more of the burden on the feature extractors, we think the number of points can be reduced to a very low level. Third, the cost of experimenting with larger parameter values did not seem justified until further results on the feature extraction are available.

4.2 The LNK Registration Procedure and Correlation

The LNK registration procedure may in some cases be viewed as a rapid procedure for computing an approximation to the cross correlation of two binary images. In this section we describe this relationship in detail and discuss its significance. Owing to the paucity of useful results on the accuracy analysis of correlation methods for registration, this relationship serves more to highlight the difficulties of analysis rather than yielding a transfer of useful theory. Despite this caveat, this relationship is important in the comparative evaluation of the LNK procedure and other registration methods.

The view of the LNK procedure as a correlation method may be clearly seen in the following situation. Assume we are given two binary images, I1 and I2, for which seek a "good" translation mapping I1 into I2. Without further information concerning the images, we can have no notion of a correct transformation; however, a meaningful comparison of the results of the LNK procedure with correlation methods may be made if we are only interested in the similarity of their conclusions.

In the restricted problem under consideration, the LNK procedure need only examine translations rather than general rigid motions. No translation will be considered unless it maps at least one abstract vector in I1 approximately into an abstract vector in I2. If the match were required to be exact, then the cross-correlation function of the two images at the lag given by this translation would have a value of at least two, since the image values are one or zero. If we require the exact matching of vectors to determine the number of matches for a given transformation, then the connection between the two procedures may be seen immediately.

The LNK procedure computes the number of matching vectors in the two images and thus implicitly computes the number of

matching points. To calculate the number of matching points, merely total the number of distinct points in I1 which are endpoints of matched vectors. Conversely, given the set of matching points, we can determine the number of matching vectors. Since the cross-correlation of two binary images represents the number of matching points at a given lag, the only difference between the procedures is that one counts matching points and one counts matching vectors. If vectors are chosen so that no point is an endpoint of more than one vector, then the two procedures yield identical results for this translation mapping of two binary images.

The assumptions made in comparing the outcomes of the two procedures bear closer scrutiny. Clearly, in remote sensing applications some tolerance must be allowed in the matching. Although the definitions of the tolerances in the LNK procedure are given in [Stockman 1980], they are somewhat complex and evolved from the need to match real vectors (edges) as well as abstract vectors. Nevertheless, the match weight computed for a pair of vectors is closely related to the matching of their endpoints. The total match weight for a given transformation may be viewed, in the case where no two vectors have a common endpoint, as the sums of the merits of point matches between the two images.

Cross-correlation is likely to perform poorly if points are sparsely distributed. A variety of methods may be employed to circumvent this difficulty. First, the image may be redefined using a coarser grid and assigning the value one to any block which contains a pixel in the original image with value one. Second, the original image may be redefined by assigning the value one to all pixels within a fixed distance of a pixel with value one in the original image. By varying the block size or fixed distance, the tolerances can be adjusted.

The effect of these two image modification procedures on the

correlation function is difficult to analyse exactly. Under the first method only the location of the block containing a point is relevant, not its location within the block. Thus if each block is d units on a side, a point P in I1 may map into a location in I2 which is as great as 2d units from a point Q in I2 and conribute as much to the correlation function as if P mapped exactly into Q. Thus this method may be viewed as a type of thresholded correlation based roughly on distance. The term threshold applies since two points conribute one to the correlation function if they are close and zero otherwise.

The second image modification procedure assigns to each pair of approximately corresponding points a measure of their closeness, namely the area of intersection of two discs, one surrounding each point. This measure is approximately monotonic as a function of the distance between points, but this may vary slightly with the distance measure chosen. In summary, the first method provides a counting measure of matches while the second gives a sum of merits from the matches.

Both the LNK procedure and cross-correlation using tolerances compute, for each translation for which there are at least two matches, the number of matches and the sum of the merits of the matches. For all other translations, the LNK procedure implicitly assigns a value of zero.

The analogy between the LNK procedure and correlation of binary images is used in Section 4.3 to compute characteristics of the LNK procedure using correlation analysis methods. We think this type of analysis may be extended considerably using further results from the correlation literature. The labelling of points may also be included in a correlation model, but we have not yet developed the point of view adequately to allow analysis of registration error.

4.3 Registration Accuracy

As discussed in Section 4.2, in the case of translation matching of binary images, the L.N.K. registration algorithm provides results similar to those obtained using the cross-correlation function. To gain a deeper understanding of the procedure, a theoretical analysis is desirable. The probabilistic analysis of the LNK procedure requires the evaluation of certain highly complex combinatorial expressions, which we are not currently able to evaluate. In consequence, we have chosen to view the LNK technique as a rapid procedure for computing the cross-correlation of the binary pictures or of a picture and a map, at least for the present analysis.

Although the remote sensing literature contains many papers analyzing correlation as a registration technique [Mostafavi 1979], the results have not been wholly satisfactory. A crucial part of these analyses, of course, is the underlying model of the image. In most works images are modelled as an array of Gaussian processes. The measures of correlation performance which are studied using this type of model have some desirable properties, and they leave many important questions unanswered.

The basic registration model assumes the presence of two images, a reference image and a sensed image:

$$\text{reference image} = Ir(x) = P(x) \quad x \in M$$
$$\text{sensor image} = Is(x) = Pd(x) + N(x)$$

Here M denotes the set of points in the image plane and $x = (x_1, x_2)$ represents a point in this plane. $P(x)$, $Pd(x)$, and $N(x)$ represent random variables where $P(x)$ gives the value of the reference image at point x, $Pd(x)$ is the contribution to the pixel value in the sensed image due to the fact that $p(x)$ has been distorted by a translation and $N(x)$ represents the contribution to the sensed image due to noise.

The image model described above assumes each image is one element of a set of possible images forming the ensemble of a stochastic process. We assume that Pd(x) is given by

$$Pd(x) = P(x+t)$$

where t is a vector representing a translation.

The basic quantity of interest in this analysis is the cross-correlation function:

$$C(t)=\Sigma Is(x-t)Ir(x)$$

where the sum is taken over the picture which has w pixels. Note that for fixed t, C(t) is a random variable since Is and Ir are random processes.

One goal of registration analysis is to compute the probability of false registration. This is easy to define in the case of most correlation procedures since in these situations a small area in a reference image is being correlated with windows in the sensed image. A false registration is a peak in the correlation function indicating a match between a segment of the reference image and a noncorresponding window in the sensed image. Because, in the discrete case, we are working with both images in their entirety, this definition must be modified. Clearly a registration in which reference image points are overlaid only on noise points from the sensed image should be regarded as a false registration. But how should we regard a registration in which several points in the reference image overlay reference points in the sensed image but not correct points. As a trivial example, if the reference image consists of four points lying at the corners of a square, then there are four ways in which we get a perfect overlay, but only one is correct.

we estimate some bounds on this type of error using simulation. In practice, any gross registration error will arise primarily from the map matching with noise. Essentially, this is a restatement of the assertion that for very sparse point images the autocorrelation tends to be very sharply peaked and have no large secondary peaks.

For purposes of analysis we assume the autocorrelation of the reference image takes on only three values: n, one, and zero, where n is the number of points in the reference image. We justify this assumption by observing that in 1,000 sets of reference image each containing five uniformly distributed points, this assertion was justified over 500 times. To appreciate the significance of this 50 percent figure, we must consider how the reference image or map is formed.

The image or map from which the reference image is formed will, in general, contain far more points than the reference image. As an example, suppose we have a map containing forty points, any of which may be selected for the reference image. If we wish to select five points, there are 658,008 such subsets. If all these subsets were independent, then, since each subset selected would have a 50 percent chance of satisfying our autocorrelation property, the probability of not finding a suitable subset in seven independent selections of five points is less than 1 percent. Though the subsets are not independent, seven disjoint sets of five points could be selected from the forty points.

The above arguments could be developed formally to get better estimates for the probability of selecting a good reference image, but we think that the assumption of the peaking of the autocorrelative function of the reference image is clearly reasonable. In practice, the selection of reference images with sharply peaked autocorrelation functions is a simple matter. This image is then a map with a single large peak in the

autocorrelation function. We need only select a small subset of map points which has the property that no two interpoint distances are close. This image is then a map with a single large peak in the autocorrelation function.

Recall that this digression on the autocorrelation function of the reference image was introduced to clarify the notion of a false registration. In summary, we have concluded that the total conrtribution to the cross-correlation function far from the true transformation due to the autocorrelation of its reference image is no greater than one. Thus we are justified in assuming that a false registration arises almost entirely from the interaction of the points in the reference image with the noise points in the sensed image.

At the true transformation the reference points should align with the corresponding image points so the number of aligned points should be n, the number of reference image points. We define a false registration to be a peak in the correlation function, greater than the value of the peak at the true transformation, located more than a few degrees and a few pixels from the peak corresponding to the correct transformation such that more than n points align at this incorrect peak. As mentioned before, the exact number of pixels and degrees remains to be determined, but five degrees and ten pixels appears reasonable based on experimental data.

We now turn to the study of C(t) where it is far from the current transformation. For such values of t, we can write our correlation function in the following form:

$$C(t)=CN(t)= \Sigma \quad N(x-t)P(x)$$

For a fixed reference image P(x) we can estimate the expected value of C(t), where the expectation is taken over an ensemble of sensed images. In this ensemble only the noisy points vary since P(x) is fixed and the sensed image contains a copy of P(x).

The probability distribution of the random variable is difficult to compute so we must make an assumption to estimate properties of the distribution. First we assume that the probability distribution of $CN(t)$ is independent of t. This is not correct since the area of overlap between the reference image and the translated noise image varies. As the overlap shrinks to zero, the probability that $C(t) = 0$ goes to one. To bound the registration error, we take the opposite extreme and assume complete overlap between the noise and the image. This assumption greatly increases the probability of false registration so the real probability of false registration is much lower than that given here.

The usefulness of probabilistic properties of $C(t)$ requires some comment. At the correct transformation, t, each reference image point lines up with sensed image point s so $C(t_0) = n$ where n is the number of points in the reference image. Far from t, at most a single reference point can line up with the image of a reference point. Thus we have

$$C(t) = CN(t)$$

or

$$C(t) = CN(t)+1$$

depending on whether or not a reference point lines up. A false registration occurs if

$$C(t_0) \leq C(t)$$

for some t far from 10. Thus a false registration cannot occur unless for some t far from t_0,

$$CN(t) > n-1$$

In light of the above remarks, we see that an upper bound on the probability of false registration is given by the probability of the following event:

E: Max $CN(t) > n-1$

We know of no way to effectively compute the probability of this event. In the analysis of correlation models for Gaussian images, a similar problem arises and the approach taken is to assume that the probability distribution of $C(t)$, for fixed t, can be used to estimate this probability of the event $\dot{E}$. This formulation assumes that the values of $C(t)$ for various t's in a given picture can be thought of as samples generated according to the distribution of $C(t)$ for fixed t but varying over the ensemble of noisy pictures.

This type of hypothesis, referred to as an ergodic hypothesis, is common in image processing models. If we further assume a functional form for the distribution of $C(t)$, then we need only estimate the parameters of the functional form to estimate the probability of false registration. In the Gaussian image case, $C(t)$ is assumed to be Gaussian, so only the first and second moments need be estimated. For the discrete case we can bypass the assumptions of a functional form and directly compute the probability, q, of the following event:

F: $CN(t) > n-1$

If w is the number of pixels in the image, then the probability of false registration, f, is given by $(1-q)^W$.

The computation of the probability q is straightforward. We assume the reference contains n points. Assume the noisy image contains m points. Recall that $CN(t)$ is a random variable whose ensemble is formed in the following way. Generate a noisy picture and shift it t with respect to the reference image. Then count the number of pixels which contain both a reference point and a noise point. This count is an element of the ensemble. As discussed earlier, to obtain an upper bound on the probability of false registration we omit the translation and assume the two images overlap completely.

Referring to each pixel as a bin, we may state our task as a

simple discrete probability problem. We are given w bins, n of which are occupied. We then take m objects and place each one in a different bin, selecting the bins using a uniform random number generator. The probability we are interested in, q, is the probability that at least $n-1$ bins contains both one of the original n objects and one of the m objects. This is given by

$$\binom{n}{n-1}\binom{w-n}{m-(n-1)} + \binom{n}{n}\binom{w-n}{m-n}$$

For a 100×100 image with five reference points and a sensed image with 25 noisy points, we have $q = 1.5 \times 10^{-10}$. In this case, F, the probability of false registration, is .0003.

To make our model more realistic we may assume the reference image contains noise points. By noise points we mean points which represent ground features but which fail to show up in the sensed image. Keeping the notation of the previous problem, we now let k denote the number of noise points in the reference image. With k noise points, $C(t)$ can be as large as min $(n+k, m)$. We shall assume $m > n+k$. The correlation of the current transformation can now range from n to $n+k$. We shall assume $C(t_0) = n$, resulting in our calculations, providing an even greater overestimate of the probability of false registration.

We are still interested in computing the probability q, but now the number of bins occupied has risen from n to $n+k$. Thus q is given by

$$\sum_{r=n-1}^{n+k} \binom{n+k}{r}\binom{w-(n+k)}{m-r}$$

For n, m, and w, as in the previous example, and $k = 5$, we obtain $q = 1.2 \times 10^{-6}$. For this value of q we have $f = .0122$. The above examples show the extremely high reliability of discrete correlation under the stated assumptions. A 100×100 image was selected to allow for a coarse resampling of a typical 512 image as suggested in Section 4.2. Thus approximate matching in the finer image is replaced by exact matching in the coarser image.

An analysis of the expected effect of this coarse sampling on the cross-correlation of the reference image with the non-noisy position of the sensed image would be useful since a coarse sampling could reduce $C(t_0)$ below n.

4.4 Sparse Correlation

Various aspects of the relationship between the LNK registration procedure and feature point correlation are discussed in Section 4.2. An analytical model of feature point correlation is given in Section 4.3. This section describes several experiments designed to estimate the registration error in discrete correlation.

Simulations were performed on one- and two-dimensional data. The image was represented as a vector of length 100 in the one-dimensional case and a 64 x 64 array in the two-dimensional case. A set of points, called map points, was then distributed into the pixels using a uniform random number generator. A copy, called the image, of the vector or array was made and additional points called noise points were thrown into these bins. The cross-correlation between the image and the map was then computed. For each such correlation the value of the cross-correlation function at the peak was recorded in addition to the correctness of the offset corresponding to this peak.

The problem of multiple points in bins was handled slightly differently in the one- and two-dimensional cases. In the one-dimensional case a bin assumed the value of one if one or more points landed in it. In the two-dimensional case the value of the bin was the number of points contained in it. The former procedure is preferable since it gives a better approximation to the number of matches. The minimum of the numbers of points in the map and image bins would actually be the best measure of how well two bins coincide, but for speed we shoose the coarser method in the two-dimensional case. We think the differences between these schemes are not sufficient to cause great differences in correlation performance, but this question could be resolved with a modest analytical and simulation effort.

The results for the one-dimensional correlation are given in Table 4.1. Note that in all cases the peak of the

Table 4.1 . One-dimensional unnormalized correlation on a binary picture of length 100 with M common points and N noisy points present. The correct transformation was always at the highest peak.

M	N	2d Highest Peak	Highest Peak
5	1	2	5
5	3	3	5
5	5	4	5
5	7	4	5
5	9	4	5
10	1	3	6
10	3	4	9
10	5	4	9
10	7	5	10
10	9	6	10
15	1	5	14
15	3	7	16
15	5	7	12
15	1	1	13
15	9	8	15
20	1	8	16
20	3	8	18
20	5	10	18
20	7	10	18
20	9	10	19

cross-correlation corresponds to the correct transformation. One purpose in investigating the one-dimensional case was to explore the possibility that the evaluation of a transformation in the LNK procedure could be accelerated by performing the matching on a horizontal or vertical projection of the points rather than working with the points in two spaces. Since the registrations were correct in all cases, we have no indication as to the limitations of correlation in one dimension.

The results of the two-dimensional correlation study are shown in Table 4.2. Matching was remarkably accurate. With as few as three common points, several hundred noise points were present before an incorrect transformation was selected. With five common points an incorrect transformation was never selected. An additional 500 trials were run in which the map always contained five points and the image had 25 additional noise points. In the 500 trials the current transformation was always at the correlation peak.

Table 4.2 . Two-dimensional unnormalized cross-correlation between two 64 x 64 images. The value of a pixel is the number of points lying within the pixels. M points are common to both images and N noise points are added to the second image.

CORRECT = Y if the true transformation corresponds to the peak of the correlation function; N otherwise

Max = value of correlative function at peak

M	N	Correct	Max
2	10	Y	2
2	10	Y	2
2	20	Y	2
2	20	Y	2
2	30	Y	2
2	30	Y	2
2	40	Y	2
2	40	N	2
2	40	Y	2
2	40	Y	2
2	40	Y	2
2	40	Y	2
2	50	Y	2
2	50	N	2
2	50	Y	2
2	50	N	2
2	50	N	2
2	50	Y	2

M	N	Correct	Max
2	60	Y	2
2	100	N	2
2	100	N	2
2	100	N	2
2	100	N	2
2	100	N	2
3	50	Y	3
3	50	Y	3
3	50	Y	3
3	50	Y	3
3	100	Y	3
3	100	Y	3
3	100	Y	3
3	100	Y	3
3	100	Y	3
3	100	Y	3
3	100	Y	3
3	100	Y	3
3	100	Y	3
3	100	Y	3
3	100	Y	3
3	200	Y	3
3	200	Y	3
3	400	Y	4
3	400	N	4
5	50	Y	5
5	50	Y	5
5	50	Y	6

M	N	Correct	Max
5	50	Y	5
5	50	Y	5
5	50	Y	5
5	50	Y	5
5	50	Y	5
5	50	Y	6
5	50	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	5
5	100	Y	6
5	100	Y	5

4.5 Distance Matching

The speed of the LNK registration procedure is dependent upon its ability to eliminate as many RT transformations as possible as soon as possible. The first screening is to check to see if an image vector and a map vector match in length before examining a transformation sending one into the other. For evaluating the remainder of the procedure, it is useful to have knowledge of the number of matches that are likely to be found at this first stage. This section describes work done in estimating the distribution of vector lengths for point images and the number of matches.

Ideally, we would like to know the probability distribution for the set of interpoint distances between n points uniformly distributed over a square. This could aid greatly in the selection of map vectors. By selecting those map vectors which have lengths less likely to occur in a random set of points, we may greatly reduce the number of transformations to be considered. Intuitively, it is clear that very long and very short vectors are unlikely to occur in a point image, but this observation is not adequate for designing an abstract vector selection procedure.

The selection of map vectors requires consideration of several conflicting goals. First, the number of map vectors should be small to reduce computation. Second, the number of map vectors should be large enough to find an adequate number of correct correspondences between image vectors and map vectors. Third, the number of image vectors should be as small as possible. Some sampling of feature points or abstract vectors, as described in Section 4.7, is usually required since the number of possible length comparisons made in initial screening goes up as n^2 m^2 where n and m are the number of image and map points respectively.

Various aspects of the vector selection problem are treated in other sections of this report. Section 4.8 describes several schemes for using labelling of points to select vectors, while Section 4.7 describes probabilistic considerations in randomly selecting vectors from a feasible set of vectors. The remainder of this section treats the role of length in the selection of map vectors.

Attempts to derive an analytical expression for the probability of various interpoint distances occurring or the expected number of times a range of distances occurs have not been successful. The basic problem in the analysis appears to be the tremendous amount of dependency among the interpoint distances. In general, when a new point is added to the picture, its location is determined by its distance from three points, thus making all other distances a function of these three. The probablity distribution for the distance between two points in the square can be easily calclated by treating the horizontal and vertical distances independently, but this sheds little light on the general case. There exists a large amount of literature on the subject of the distribution of random geometric objects [Santalo 1976], but we haven't yet been able to utilize any of this work.

The distribution of interpoint distances for an image containing 25 points was estimated using simulation. One hundred pictures were generated, each occupying a 512-by-512 square. The distribution of the interpoint distances is given in Table 4.3. As expected, the distribution is unimodal. This table can be used to obtain a rough estimate for the expected number of matches between a map vector of specified length and the image.

The distribution in Table 4.3 allows only a rough approxmation in computing the expected number of matches between an image and a map. A study was performed to directly estimate the number of matches based solely on edge length. The

Table 4.3 . Distribution of interpoint distances for 25 randomly distributed points in a 512 x 512 image.

Distance	Probability of Occurrence
0-40	.02
41-80	.05
81-120	.07
121-160	.09
161-200	.10
201-240	.11
241-280	.11
281-320	.11
321-360	.10
361-400	.10
401-440	.10
441-480	.08
481-520	.07
521-560	.05
561-600	.03
601-640	.01
641-680	.01
680-	.00

experiment consisted of 25 trials. In each trial a 512 x 512 map consisting of five uniformly distributed map points was created. The map was then rotated 5 degrees, with a shift of 10 pixels in the x direction and 20 in the y direction. Map points were slightly perturbed by adding random perturbations to the x and y coordinates of each point. The perturbations are generated from a Gaussian distribution of mean zero and variance one. Varying numbers of noise points were added, ranging from fifteen to sixty. All interpoint distances among the image points and among the map points were used. For each trial, each image interpoint distance which was within ten pixels of a map interpoint distance contributed one to the total number of transformations.

In each trial at least ten of the transformations arose from the correct correspondences between map edges. Table 4.4 provides a rough estimate of the size of the registration problem. Based upon the number of transformations observed in this experiment, we think that if clustering is to be eliminated, a much more selective screening process is required, such as the point labelling described in Section 4.8. These simulations allow us to form probabilistic estimates for the performance of abstract vector selection. In Section 4.7 the vector sampling problem was considered from the viewpoint of determining the probability of selection of an adequate number of correctly matching vectors. Table 4.4 provides an estimate of the total number of transformations to be considered.

Based on Table 4.4 and the assumption that very long vectors are more stable than very short ones, an apparently reasonable vector sampling scheme would be to choose only long vectors. Unfortunately, this scheme fails to take into account the relative size and overlap of map and image. If the map is much larger than the image, then the longer map vectors won't even fit in the image. Similarly, if the overlap between map and image is much less than the image size, then the maximal length of a matching vector is reduced. If the image and map are both square

Table 4.4 . Number of image edges which agree in length within ten of a map edge. The maps contain five randomly generated points and the image contains these five plus N additional noise points.

N	# of Matches	N	# of Matches
15	116	40	418
15	92	40	483
15	123	40	454
15	73	40	485
20	142	50	600
20	160	50	750
20	135	50	678
20	159	50	703
20	145	50	631
30	274		
30	222		
30	260		
30	316		
30	256		

and one is horizontally shifted half an image length with respect to the other, then the lenth of a maximal vector is reduced by 20 percent.

Summing up the merits of various types of vectors, we may make the following remarks. Short vectors are good because they tend to have fewer spurious matches but bad because small changes in their endpoints can lead to large changes in rotation angle. Medium-length vectors are better than short vectors with regard to the rotation problem but worse with regard to the number of spurious matches and the possibilities of lying outside the image-map intersection. Finally, long vectors are most insensitive to endpoint changes, unlikely to produce many false matches but more likely to not lie in the image-map intersection.

We would like to quantify the above ideas so a merit could be assigned to each vector. Selection of vectors could then be based on this merit. Two of the three factors can be easily quantified. The effect of vector length on rotation angle could be easily simulated. The number of possible matches can be estimated using Table 4.4 and similar computations with varying numbers of points. Simulations can also be performed to estimate the effect of vector length on the vector lying outside the image map overlap. Assumptions on the type of overlap between image and map may affect this estimation considerably. If the only assumption on overlap is the approximate amount of common area, the results may vary greatly with orientation. We have just begun study on these questions.

A further problem in the evaluation of vectors arises if the merits canot be considered independently. There are at least two ways in which we may have interaction between the selected vectors. First, points shared between vectors result in fewer points being represented by the vectors. Second, the distribution of vectors in the image is important. If most vectors selected are in one part of the image, we may fail to get

enough vectors in the map-image intersection. Other types of dependence may be present. In spite of these dependencies, we think that the independence assumption may be reasonable as a starting point for defining a merit function for vectors.

4.0 Empirical Registration Accuracy

This section describes the results of simulation studies in which the LNK procedure was applied to abstract vectors formed from randomly generated point images. We were interested in analysing the deterioration in registration accuracy as the number of non-corresponding points between images is increased. The results of this section served to bring out the need for the study of sampling and labelling procedures as described in Section 4 of this report.

The artificial images used in this study were of two types. First, a point image, called a map and consisting of five to ten points was generated, where a uniform random number generator was used to generate the x and y coordinates of the points. Then a perturbed copy, called the image, of this map was formed by rotating the image five degrees, and translating it ten units in the x direction and twenty units in the y direction. In addition, the x and y coordinates of each point were perturbed by adding independent random numbers to the x and y coordinates where the random numbers were generated by a zero mean Gaussian process, with varying variances. Then a set of additional points, called noise points was placed in the image by uniformly generating x and y coordinates.

The LNK registration procedure requires abstract vectors to be formed from the feature points so two simple schemes were selected. In one scheme all pairs of points were selected as abstract vectors while in the other scheme pairs of points were randomly selected. For each pair of points chosen, one was randomly chosen to be the head of the vector and the other to be the tail.

The match weight of a transformation was the measure of registration quality used for determining the registration transformation. It weighs both the average separation between

the vectors and their lengths. Though this measure of registration quality is not ideal, it provided sufficient insight into registration accuracy for the present study. For further information on match weights, see [Stockman 1981].

The results of the simulations are given in Table 4.5. The success of a given trial is determined by looking at the θ, x, and y columns of the line in the table corresponding to the trial. For a perfect registration, these variables should assume the values 5, 10, and 20 respectively. These variables represent the rotation and translation corresponding to the highest match weight transformation examined by the clustering program, where the weight is given in column W_TP. The match weight of the correct transformation is given in Column WTT. In cases 7, 8, 38, 39, 41, 42, and 46, the highest match weight transformation considered by the clustering routine was far from the correct transformation, but the match weight at this transformation was significantly lower than the match weight at the correct transformation. For several of these cases detailed examination of the program output led to the conclusion that the correct transformation was not near any of the clusters examined by the clustering program.

The primary problem with the clustering program appears to be in the way in which bins are formed in the search for clusters. To guarantee the detection of cluster peaks, bins must be overlapped. Otherwise, a cluster can be spaced over several bins without enough of a concentration in any single bin to be singled out as a potential cluster center. The cost of providing enough bins to prevent this overlap problem can be high, especially when scaling is added to the problem. The consideration of more clusters could be obtained by lowering the standards for a cluster peak, but this also adds considerable cost to the algorithm.

In addition to the registration failures mentioned above, incorrect registrations were obtained in cases 11, 12, 32, 33, 34, 35, 37, 44, 45, and 47. The match weight at the correct

transformation in case 11 was 1990 while the next highest matchweight at the correct transformation was 917. For those cases in which an approximately correct transformation was predicted by the program, the lowest match weight at the correct transformation was 1246. Thus registration failures appeared to be arising from two problems. First, the clustering routine may fail to label the correct transformation as being near a cluster peak, even though the match weight at that transformation is high. Second, the match weight at the true transformation may be low, resulting in failure to recognize this as a cluster peak.

The significance of the error sources requires further comment. Only in 1 case out of 47 did an error occur in which the match weight at the correct transformation was substantial, say above 1000, and the match weight at the best transformation predicted by the algorithm was larger. Apart from errors due to the clustering, it appears that errors are due more to fewer correct vectors matching than to too many incorrect vectors matching.

Further corroboration of the hypothesis that a lack of good matches rather than too many bad matches creates errors may be found by looking at the match weights at the predicted registration transformations in the cases where an error occurred which was not due to clustering. Excluding case 11, because its match weight at the correct transformation was high, the highest match weight in the remaining such error cases was 1521 at the predicted transformation. This was higher than the correct match weight at the true transformation for only five of the thirty correct registration cases.

By setting a minimum required match weight of 2000 to accept a predicted transformation, we obtain 25 successful registrations, 1 incorrect registration, and 26 registrations rejected due to inadequate match weight. If we had instead always used the predicted transformation without regard for a minimum match weight, then 30 correct registrations and 17 incorrect registrations would have been obtained. these figures are used to indicate the level of improvement available in

registration if we are willing to reject registrations. The actual percentages of correct and incorrect registrations do not reflect on the algorithm since we chose values for the number of map and image points to provide data on how the algorithm behaves near parameter values where it breaks down.

The tradeoff between algorithm accuracy and the number of noise and map points present is difficult to describe with precision. Using the simple random selection of edges, the program generally produced correct results as long as the number of noise points was no more than 2.5 times the number of map points. This is a very rough guideline which we did not feel merited further empirical examination until a better analytical understanding of transformation screening was prformed. As a result of the studies described in Sections 4.7 - 4.11, we feel that this factor of 2.5 can be improved considerably.

The cases in which registration failure occurred are due primarily to an inadequate number of correct image vectors being given to the registration algorithm. Our studies in Sections 4.7 - 4.11 indicate that a much larger number of correct image vectors can be given to the registration program if more pre-screening is performed, including point labelling and matching of triples of points. The former procedures require refinements to the feature selection process while the latter can be used currently.

The cases studied in this section provided considerably more data than is given in Table 4.5. For each case the match weights are given for several cluster peaks for several levels of clustering. This more detailed information will prove useful in setting reasonable criteria for accepting or rejecting a registration and for more detailed study of clustering if it is to be retained in the algorithm.

We have dealt only with the factors involved in getting an approximately correct registration. A treatment of registration accuracy given that an approximate registration has been obtained involves very difficult methods of analysis. We did not think it reasonable to examine this question until a better understanding

of approximate accuracy was obtained.

In summary, the LNK procedure is very robust with respect to the number of erroneous image vectors as long as an adequate number of correct vectors are present. Using somewhat loose terminology, it is far more important in obtaining a good registration to have a correct signal having at least a certain minimum strength than it is to have a high signal-to-noise ratio.

Table 4.5 Registration results using LNK procedure on a map with CP points and an image with CP+NP points where the correct transformation is =5, x=10, y=20. The image contains a perturbed copy of the map with variance V (see text) and NP additional noise points. IE and ME are the number of image and map vectors used. A denotes all vectors. $_wT_T$ is the match weight at the correct transformation and WTP is the the predicted transformation parameters. The remaining columns are not relevant at present.

C	CP	NP	V	IE	ME	WTT	TI	TI		X	Y	WTP	PI	PM
1	8	17	5	100	A	3603	11/100	11/28	4	9	19	3389	10/100	10/28
2	8	17	1	100	A	2807	7/100	8/28	6	9	19	292	7/100	8/28
3	8	17	5	100	A	2610	8/100	8/28	6	9	19	2807	8/100	8/28
4	8	17	1	100	A	1503	5/150	5/28	4	11	19	1514	5/100	5/28
5	8	15	1	100	A	1442	5/100	5/28	4	9	24	1432	5/100	5/28
6	8	15	1	100	A	2482	7/100	7/28	4	9	17	2442	7/100	8/28
7	8	20	1	100	A	1739	6/100	5/28	150	86	1	1264	5/100	5/28
8	8	20	1	100	A	1225	6/100	5/28	128	50	87	646	2/100	2/28
9	8	20	1	100	A	1367	7/100	7/28	6	9	11	1771	7/100	7/28
10	8	15	1	100	A	3132	9/100	9/28	4	3	15	2946	9/100	9/28
11	5	20	1	100	A	1990	2/100	2/10	184	218	17	2910	3/100	33/10
12	5	20	1	50	A	0	0/50	0/10	243	46	228	980	1/5	1/10
13	5	0	5	10	A	5850	6/10	6/10	3	14	15	5930	6/10	6/10
14	5	10	5	A	A	9770	10/105	10/10	4	14	13	9590	10/105	10/10
15	5	7	3	A	A	9770	10/66	10/10	4	14	13	9590	10/66	10/10
16	5	0	0	A	A	8890	9/10	9/10	4	7	18	8890	9/10	9/10
17	5	5	10	A	A	8890	11/45	9/10	4	7	18	8880	11/45	9/10
18	5	5	10	A	A	6900	8/45	7/10	3	15	12	6950	8/45	7/10
19	5	5	7	A	A	6920	7/45	7/10	3	15	13	6930	7/45	7/10
20	5	5	5	A	A	9770	10/45	10/10	4	14	13	9590	10/45	10/10
21	5	15	5	A	A	9770	11/190	10/10	4	13	14	9700	11/190	10/10
22	10	10	10	100	30	4246	13/100	13/30	6	7	19	4220	13/100	13/30
23	10	10	10	100	30	6136	22/100	14/30	4	11	19	6160	22/100	19/30
24	10	10	15	100	30	5626	18/100	18/30	6	19	17	5840	14/100	18/30
25	10	10	15	100	30	5140	18/200	16/30	6	11	27	5130	18/100	16/30
26	10	10	20	100	30	4303	14/100	14/30	6	11	16	4346	14/100	14/30

27	10	10	20	100	30	4870	16/100	15/30	6	9	19	4876	16/100	15/30
28	10	10	25	100	30	4833	16/100	16/30	4	9	19	4833	16/100	15/30
29	10	10	25	100	30	5056	16/100	16/30	4	7	17	4750	17/100	15/30
30	10	10	30	100	30	3850	11/100	12/30	4	13	16	3856	11/100	12/30
31	10	10	30	100	30	4056	12/100	13/30	4	7	24	4163	12/100	13/30
32	8	35	5	100	A	9	0/100	0/28	301	-47	19	939	3/100	3/28
33	8	35	5	100	A	0	0/100	0/28	56	-240	5	960	2/100	3/28
34	8	30	5	100	A	325	1/100	1/28	6	-78	-95	1521	6/100	6/28
35	8	30	5	100	A	357	1/100	1/28	118	-119	47	1100	4/100	4/28
36	8	30	5	100	A	1246	5/100	5/28	4	17	21	1275	5/100	5/28
37	8	30	5	100	A	0	0/100	0/28	69	-15	85	1196	5/100	5/28
38	8	30	5	100	A	1000	3/100	3/28	237	-70	207	675	2/100	2/28
39	8	25	5	100	A	2010	6/100	6/28	237	54	-113	1539	6/100	5/28
40	8	25	5	100	A	2103	6/100	6/28	4	11	20	2117	6/100	6/28
41	8	35	5	100	A	1160	4/100	4/28	297	110	209	971	3/100	3/28
42	8	35	5	100	A	1035	4/100	4/28	193	-23	-96	846	4/100	3/28
43	8	35	5	100	A	1435	5/100	5/28	6	3	19	1635	6/100	6/28
44	8	25	5	100	A	735	4/100	3/28	1	61	134	1128	4/100	4/28
45	8	25	5	100	A	917	3/100	3/28	325	-23	56	1228	4/100	4/28
46	8	25	5	100	A	1746	4/100	5/28	324	5	59	1017	2/100	3/28
47	8	25	5	100	A	735	3/100	3/28	171	97	5	778	3/100	3/28

4.7 Vector Sampling for Registration

The LNK registration procedure does not currently provide a method for the selection of abstract vectors from feature points. As shown in Section 4.8, labelling of point features can lead to a significant reduction in the number of possible abstract vectors. In spite of this reduction, it may not be feasible to examine the full set of transformations determined by all pairs of matching vectors. Several alternatives to the examination of all transformations are possible. First, a clustering procedure can be applied to the set of transformations in order to locate transformations which are supported by many matching vectors. Second, a subset of vectors may be randomly selected and all possible transformtions determined by those vectors may be examined. This section treats some of the combinatorial problems associated with these schemes in order to aid in the design of procedures for selecting transformations.

The clustering procedure currently used in the LNK scheme has several severe limitations, described in Section 5. Apart from the possibility of totally missing the correct cluster, as occurred in several experiments, the clustering procedure is slow enough to cause a bottleneck in the registration process if a large number of abstract vectors are used. The experiments described in Section 4.5 show that the number of feasible transformations can grow quite rapidly with increases in the number of noise points. The clustering of this many transformations puts a heavy computational burden on the registration process.

The empirical determination of the histogram of interpoint distances for randomly distributed points shows that the lengths appear to have a unimodal distribution. Thus, if a small set of points is selected from the map, the vector lengths will tend to be clumped around the mean and many image vectors will tend to match. On the other hand, if very long or very short image

vectors are selected, very few matches are likely to be present. Further experimentation should be done to determine the reduction in the number of matches when long vectors are selected.

A variety of deterministic schemes can be used to limit the number of vectors to be used for matching. These procedures, including selection of long edges and labelling of point features, may reduce the number of vectors significantly, but a large number of vectors may remain even after this screening. In addition, the screening may be expensive, especially in the case of point feature labelling. If the number of vectors remaining is still too large to use, some random sampling of the vectors must be made. Alternating random sampling of the points could be done.

Each procedure for screening vectors leads to a different sampling problem. Only a few of the problms have been examined to date. The basic goal of our sampling procedure should be to obtain a set of vectors adequate for an accurate registration. Since an answer to the question would require a deeper analysis of registration accuracy than is currently available, we treat a simpler problem: how small a random sample of image vectors can be selected which will, with high probability, contain at least a specified number of map vectors, i.e. image vectors corresponding to map vectors?

We now describe the sample problem for a map containing f points and an image containing v points, where the map points are also in the image. The number of map vectors is f(f-1), since any map point may be chosen as the tail of a map vector and any of the remaining points may be chosen as the head. Similarly, the image contains v(v-1) vectors. Many of these image vectors will not match with any map vectors. Let p be the number of image vectors matching with at least one map vector. If all map vectors are used, then k of the p vectors represent correct matches.

Several sampling procedures can be applied at this point. The simplest is to select a fixed number of the p feasible image vectors. A more sophisticated procedure would be to associate to each map vector the set of image vectors associated with it and to sample from each of these sets. The latter procedure appears promising, for if several of the sets are very small, they may be sampled exhaustively, thereby guaranteeing that a lower bound on the number of correct image vectors present is the number of small sets.

Both of the above sampling procedures are analytically tractable. For the first procedure we randomly select a subset containing n of the p vectors. Assume we wish to choose n so that one set will contain at least s map vectors. The probability of getting exactly t map vectors for any 0 t m is easily calculated by observing that a sample of r vectors containing t map vectors contains r − t image vectors which are not also in the map. Since there are a total of m map vectors, the t map vectors can be chosen in $\binom{m}{t}$ ways. Similary, the non-map vectors can be chosen in $\binom{p-m}{r-t}$ ways. As there are $\binom{p}{r}$ ways to choose the r vectors, the probability of picking exactly t map vectors is given by

$$p(x=t) = \frac{\binom{m}{t}\binom{p-m}{r-t}}{\binom{p}{r}}$$

where p(x=t) denotes the probability of picking exactly t map vectors.

The probability density p(x=t) is just the probability density function for the hypergeometric distribution with parameter r, m, and p [Johnson and Katz, 1969]. To denote the explicit dependence of p(x=t) on the parameters of the distribution, we denote p(x=t) by $p(t|\,r,m,p)$. Since we wish to obtain at leases map vectors, we should choose r so that

$$H(s;r,m,p) = \sum_{t=s}^{b} p(t\ r,m,p) > \alpha$$

where b = min (m,r) and is the desired confidence level and H(s;r,m,p) is the probability of obtaining at least s map vectors. Noting that the same hypergeometric distribution is obtained if r and m are interchanged, a number of approximation procedures for H(s;r,m,p) may be applied.

The quality of the approximations to H(s;r,m,p) depends upon the relationship between the parameters of the distribution. Before describing these approximations, we stress that the point in evaluating H(s;r,m,p) is to find the smallest r such that H(s;r,m,p) > α. Since approximations are normaly given in terms of H(s;r,m,p) = 1-H(s;r,m,p), our problem is to find the smallest r such that H(s;r,m,p)<1 $-\alpha$. We know of no analytical procedure to find the smallest suitable r. Using the fact that H is a monotonically increasing function of r, a simple binary search for the smallest r is possible.

We do not yet have computational experience with approximations to the hypergeometric function, so we cannot make recommendations as to a suitable choice at this time. The following approximations appear potentially relevant [Johnson and Kotz 1969]. For m > r, which can always be obtained (see note following the definition of H), the following approximation has been constructed:

$$\bar{H}\ (s) = \sum_{t=0}^{s} p(t|r,m,p) \quad \sum_{t=0}^{s} \binom{r}{t}(m/p)^t\ (1-(M/p))^{r-t}$$

$$-\frac{1}{2}\,,\ \frac{r(r-1)(m/p)(1-m/p)}{p-1} \quad \binom{r}{s+2}\ (m/p)^{s+2}(1-m/p)^{r-s-2}$$

$$-\binom{r}{s+1}(m/p)^{s+1}(1-m/p)^{r-s-1}$$

A second approximation is given by

$$\bar{\mu}(s) = \sum_{t=0}^{s} \binom{r}{t} w^{j}(1-w)^{r-j}$$

where $w = (m-1/2s)/(p-1/2r + t1/2)$.

When more understanding of the range of parameter values occurring in real data has been obtained, we can make a decision as to whether r sould be computed or determined by a table lookup.

In the other sampling procedure to be considered, each map vector has associated with it a set of image vectors of similar lengths and each set is sampled. It is possible for an image vector to match two map vectors, in which case the sets are not disjoint. To avoid this problem, we assume that no two map edges are similar in length. See Section 4.5 for a discussion of the feasibility of such a selection.

Our basic sampling problem may be stated in the following way: We are given d sets $S_1, . . . , S_d$ of image vectors which form a partition of the set of feasible image vectors. Each set is guaranteed to contain exactly one map vector. Select the smallest number of samples r such that, with a probability of at least α, we will have s map edges in the sample.

The selection of vectors from each set is now analogous to our previous sampling procedure except that the number of map vectors selected can only be zero or one. One method is to select the s smallest sets where s is the number of map vectors we wish to obtain and select a sufficiently large set of samples so that the probability of selecting map vectors from all S_i's under consideration is greater than α. As far as a single set S_i is concerned, the map contains only one vector, and we wish to guarantee with given probablity that the vector is selected. The probablity of selecting exactly one map vector in a sample of size r is given by

$$p(x=1) = \frac{\binom{1}{1}\binom{n_i-1}{r_i-1}}{\binom{n_i}{r_i}} = \frac{r_i}{n_i}$$

Assume, without loss of generality, that the first s sets, $S_1, \ldots, S_s$ are the smallest sets. Then the probability of obtaining a map edge in each sample is given by

$$\prod_{i=1}^{s} (r_i/n_i)$$

We wish to select the r 's so as to minimize

$$\sum_{i=1}^{s} r_i$$

subject to the constraint that

$$\prod_{i=1}^{s} (r_i/n_i) > \alpha$$

We are not aware of a rapid procedure for solving this problem, so if we constrain ourselves further to sample the same fraction of each set so r_i/n_i is the same for all i, and this quantity is denoted by d, we have $d^s > \alpha$, so $d = \alpha^{1/s}$. For each i, define r_i to be the smallest integer greater than or equal to $n_i d$.

Other sampling schemes are possible to handle the sampling from the S_i 's, though we have not investigated them. The sampling in the case of the sets clearly requires a very large fraction of each set so it may be best to just use all of the s smallest S_i's. If the number of vectors in the s smallest S_i 's

is very large, the sampling results can be used to set a threshold above which registration is likely to fail, so more sophisticated procedures such as feature point labelling should be attempted.

One limitation on the above sampling procedures is based on the assumption that a lower bound on the number of map vectors found in the image should be determined. A more useful bound would be on the number of map points which have been matched. Matching map points rather than vectors is more meaningful since the vectors are generated from the points. To appreciate the potential difference between the number of matched vectors and the number of matched points, we consider two extreme cases. If we have n map vectors which have to be mapped and all have an endpoint in common, then only n - 1 points have been matched. If no two vectors have an endpoint in common, then n points have been matched.

Further study should be done to determine the relationship between the number of matched vectors and matched points. Sampling procedures which give information on the distribution of the number of matched points can be defined, but extensive analysis has not yet been completed. A simple way to reduce the problem to our previous sampling problem is to select the set of map edges so that no two edges have a point in common.

Some sampling procedures for labelled feature points can be analyzed easily. If all points are labelled as in Section 4.8, and only points with different labels are joined, then we may view the vectors as being divided into two classes: those which are also map vectors, and those which are not. We are then interested in selecting a subset which is sufficiently large to contan a spcified number of map vectors. The analysis is now identical to that in our first random vector selection procedure.

4.3 Feature Point Labelling

The speed and accuracy of the LNK registration procedure is due to the algorithm's ability to eliminate consideration of many infeasible transformations in contrast to ordinary correlation procedures. Rather than allowing all possible transformations and using the data to select the most suitable one in standard correlation computations, the set of transformations to be considered is generated by the data. Placing more structure upon the data and using this additional structure reduces the number of transformations considered and hence increases the speed and accuracy of the procedure.

A simple but effective enhancement of the LNK procedure results if each feature point has a type or label associated with it. This type may be defined in many ways. If the feature point is an intersection of lines such as roads, the number of lines entering the intersection, the angle at the intersection, and the classification of pixels in a neighborhood of the intersection may all be used to categorize this point.

The labelling of feature points can result in great speedups in the matching of abstract vectors, but this must be balanced against the cost and accuracy of the labelling. At one extreme, enough information about a point may be available to avoid any confusion of this point with any other point in the image. A matching of two such points with the corresponding points in a second image or map is adequate to determine the RT transformation, thus reducing the remainder of the abstract vector matching to a triviality.

We do not, at present, have adequate information on the cost-performance tradeoffs associated with feature point typing. Crude approaches like the Moravec interest operator or spot detectors are unlikely to be directly useful for accurate labelling [Lambird 1980]. However, they can be extremely fast

and, in some cases, may make excellent use of parallel architecture. More elaborate feature point detectors, such as the LNK intersection procedures (Section 3) may provide far more information about the point, but the cost may be considerable.

The selection of an algorithm for typing points and an appropriate modification of the LNK procedure to use this typing will require an assessment of several factors. First, how rapidly and accurately point features can be obtained. Second, how much this labelling disambiguates the points. Many points may have similar neighborhoods, so local operations may yield only a few types of points. Third, what improvements will result in the abstract vector approach as a result of various labellings. For example, if an image contains twenty feature points, what advantage results from being able to divide the twenty points into four types, each containing five points, as opposed to two groups of ten points each.

While all three of these facets of labelling must ultimately be addressed, we could not address the second issue, the extent of disambiguation possible, until the feature point selection algorithms (see Section 3) were further developed and tested. For the remainder of this section, we assume image and map feature points have been extracted and consider the effects of varying the number of types and the number of points of each type.

The LNK registrative procedure involves several major steps, each of which is affected by the typing of points. For the analysis in this section, we regard the basic steps as: 1) creation of abstract vectors, assuming point features are given; 2) selection of transformations to be clustered; 3) clustering of transformations; 4) evaluation of each clustered transformation; and 5) selection of the best transformation and an assessment of the correctness of the transformation.

Step 5) requires explanation since, in practice, it would seem reasonable to select the transformation associated with the highest match weight as the best transformation; and if this transformation leads to inconsistencies in later processing, to try to perturb it slightly to get the next-best match weight or other acceptable transformation. Although this procedure is reasonable if all point features are obtained before abstract vectors are formed, the interleaving of registration with feature point selection described in Section 7 requires an evaluation of a registration transformation based upon a small number of feature points to determine if more feature points should be sought.

The first step of the LNK procedure, the selection of abstract vectors, may be greatly affected by the typing of points. As we have seen in Section 4.7, several obvious strategies for selecting abstract vectors from feature points may be readily defined. Each of these selection strategies leads to several new selection strategies when typed points are available. Important questions in the practical analysis of certain of these strategies are unresolved, and the partial analysis of all would lead to a profusion of results without providing an adequate basis for comparison. Consequently, the present discussion will focus on the simpler procedures for which more readily interpretable results are available.

The first case we consider is the image-to-map registration problem in which the number of map points is sufficiently small so that every pair of map points defines an abstract vector. Furthermore, we assume the number of image points is sufficiently large that not all pairs of image points should be used as abstract vectors. Let n denote the number of map points, m the number of image points, and k the number of image edges selected.

The above situation was analyzed in Section 4.7, where the expected number of image edges selected which corresponded to map

edges was computed. The comparison of this result with the expected number of edges for typed matches depends on the scheme for using the typing. Several possibilities will be considered and their performances compared.

A simple abstract vector selection scheme is to connect all pairs of points of the same type by an abstract vector and to connect no points of different types. If the m image points are divided into p groups, $G_1, \ldots G_p$, where group G_i contains q_i points, then the total number of image edges selected is

$$\sum_{i=1}^{p} \binom{q_i}{2}$$

The total number of image edges is $\binom{m}{2}$ where $m = \sum_{i=1}^{p} q_i$. As a special case, assume all groups are the same size, i.e., $q_1 = q_2 = \ldots = q_p$. Then IER, defined to be the ratio of the number of image edges using typed point to the total number of image edges, is given by

$$\mathrm{IER} = \Sigma \binom{q_i}{2} \Big/ \binom{m}{2}$$
$$= p \binom{m/p}{2} \Big/ \binom{m}{2}$$

For example, if m = 30, p = 6, we have IER = 60/435, resulting in a ten percent reduction in the number of image edges to be considered. Given that p groups have been selected, it is natural to ask how IER varies as the distribution of points in groups varies.

With the aid of a simple combinatorial identity, it is easy to see that, for fixed p and m where p divides m, $\mathrm{IE_R}$ is minimized when $q_1 = q_2 = \ldots = q_p$. The identity is closely

related to the fact that the geometric mean of two numbers is less than or equal to the arithmetic mean.

Prop. For any integers a and b with $a > b$, we have the following inequality

$$\binom{a+1}{2} + \binom{b-1}{2} > \binom{a}{2} + \binom{b}{2}$$

Proof: $a > b-1$

$$2a > 2b-2$$

$$a-3b+2 > -a-b$$

$$a + a+b - 3b+2 > a - a+b - b$$

$$(a+1)a+(b-1)(b-2) > a(a-1)+b(b-1)$$

$$\frac{(a+1)!}{(a-1!2!} + \frac{(b-1)!}{(b-3)!2!} < \frac{a!}{(a-2)!2!} + \frac{b!}{(b-2)!2!}$$

$$\binom{a+1}{2} + \binom{b-1}{2} < \binom{a}{2} + \binom{b}{2}$$

If we have two groups of points with a points and b points respectively with $a > b$, and if all points within a group are joined by edges, the proposition tells us that the total number of edges increases as the difference between a and b increases. This shows that if we have only two types of points in the image, the maximum reduction in the number of edges occurs when the two groups are of equal size. The general case of p groups can be handled by induction.

The selection scheme which selects edges by joining points of similar types has been shown to perform best when all groups are of equal size and the percent improvement edge ratio has been computed. Simplifying the expression for IER in the case of p groups of equal size, we get

$$IER = \frac{(m/p-1)}{m-1}$$

For large m and p much smaller than m, we have $m/p-1 \approx m/p$ and $m/(m-1) \approx 1$, so IER $\approx$ $1/p$. Hence in this case the division of points into p groups of equal size reduces the total number of edges by a factor of p. This reduction factor is, indirectly, a measure of the speedup of the procedure resulting from labelling points. Since the speedup deteriorates as groups become more unequal in size, it is useful to find a procedure which performs better in the case of great variation in group size. To make effective use of this pair of procedures, a measure of group size equality would be desirable to help select the appropriate procedure.

The joining of points within groups results in an edge selection procedure which is effective when the groups are of equal size, so it is natural to consider whether joining points between groups might be more effective when the groups are of unequal size. The obvious analog is to join two points by an abstract vector if the points do not lie in the same group. These two approaches are complementary since every abstract edge joining two points of the image join either two points in the same group or two points in different groups. Thus a simple procedure for selecting between the two edge selection procedures is to compute the number of edges in the two cases and select the procedure which uses the smaller number of edges.

If we fix the number of groups at p, then it is easy to show that the maximum number, d, of within group edges occurs when all but one group contains exactly one point. This is easily seen to be the grouping which yields the minimum number of between group edges. The worst case performance for the two procedures occurs when the number of intergroup edges is equal to the number of intra-group edges. This occurs when each set of edges is one-half the total number of edges, i.e., the number of edges is m/2, resulting in an IER of 1/2. Thus by an appropriate choice of edge selection routines, the number of edges can be reduced by a factor ranging from 1/2 to approximately 1/p where p is the

number of groups.

Table 4.6 . Number of matching vectors between a 5-point map and an image containing the map and N noise points. Each map point has a distinct label, and those labels are randomly distributed among the noise points. Two edges match if their endpoint types match and their lengths differ by less than ten.

N	# of Matches	N	# of Matches
15	17	40	41
15	12	40	45
15	13	40	44
15	19	40	33
15	16	40	54
15	17	50	54
15	15	50	79
20	20	50	78
20	21	50	63
20	23	50	60
20	20	50	69
20	15		
30	20		
30	33		
30	29		
30	38		
30	32		

4.9 Point Labelling Experiment

Labelled feature points appear promising for reducing the number of RT transformations which must be evaluated. This section describes an experiment designed to estimate the average reduction. Within most of the range of parameters investigated, the reduction in the number of transformations is sufficient to allow all remaining transformations to be evaluated in detail, thus eliminating the need for clustering.

The experiment consisted of two phases. The first phase consisted of an empirical estimation of the average number of matches between image and map vectors. Each trial consisted of a randomly generated map containing five points and randomly generated images containing a rotated, translated perturbed of the five map points in addition to background noise points. The horizontal and vertical coordinates of all points were generated according to a uniform distribution. In copying the map points into the image, the x and y coordinates were perturbed by adding vectors generated from a bivariate gaussian distribution of mean zero and variance one. Both image and map were 512 x 512 pixels.

A count of the number of matching interpoint distances between image and map was computed. A pair of map points was declared to match a pair of image points if the vector joining the image points differed in lengths by no more than ten from the vector joining the map points. The results are given in Table 4.6.

The matching of the map with the perturbed version of the map will lead, with very high probability, to at least ten matches. Since there are ten map edges, more matches may arise if the map has non-unique edge lengths. The percentage of correct matches (approximately ten) to the total number of matches varies from about 20 percent to 1.5 percent. Based on the results in Table 4.6, exhaustive examination of

transformations determined by the labelling restriction is feasible for maps containing as many as forty noise points. It may actually be feasible for larger numbers of noise points, but more work must be done in evaluating the cost of finding matchweights before this can be determined.

This phase of the experiment provides useful information in several ways. First, it serves as a benchmark against which we may compare the reduction in matches resulting from labelling. Second, the results provide a measure of comparison in evaluating the desirability of clustering transformations versus examining all transformations corresponding to matches. Since the time required for matching vectors depends in many cases only upon the number of matches to be performed, and not on the lengths of the vectors, an estimate of the comparison time per vector can be used in conjunction with Table 4.6 to obtain an estimate of the time required by the registration algorithm. A third use of the table is in the analysis of random sampling of matching edges to speed up the registration program (see Section 4.7).

4.10 Triples for Matching

The first phase of the LNK registration procedure determines those transformations for which at least one image vector matches with a map vector. This is equivalent to determining those transformations for which a pair of map points matches a pair of image points. To reduce the number of transformations we may further restrict our attention to those transformations which result in at least three points matching approximately. This secion describes some theoretical considerations in evaluating the relative efficiency of matching triples versus pairs of points.

The average number of matching triples between a map and an image is difficult to determine analytically, though some useful results are easily described. Assume we have a map containing m points and an image containing n points including the m map points. There are $\binom{m}{2} = m(m-1)/2$ common pairs of points at the correct transformation due to all the map points overlaying. There are $\binom{m}{3} = m(m-1)(m-2)/6$ common triples of points at the correct transformation. Thus the number of triples is $(m-2)/3$ times as great as the number of pairs at the correct transformation. Similarly, for any transformation for which K points correspond, the number of matching triples is $(k-2)/3$ times the number of matching pairs. Thus if k_i transformations result in exactly i points matching, the number of matching triples is

$$T = \sum_{i=3}^{m} k_i \; i(i-1)(i-2)/6$$

while the number of matching pairs is

$$P = \sum_{i=2}^{m} k_i \; i(i - 1)/2$$

We now consider the above problem in the special case where the number of map points is five. We have

$$T = K_3 + 4k_4 + 10k_5$$
$$P = K_2 + 3k_3 + 6K_4 + 10k_5$$

The selection of triples rather than pairs of matching points is justified when $P - T$ is sufficiently large. Note that $P - T = 2k \; + 2k \; + k$, so the number of pairs is at least as large as the number of triples. More generally we have the following:

$$P - T = k_2 + \sum_{i=3}^{m} k \; i(i - 1)\,(1/2)\,[1 + (i - 2)/3].$$

Thus the number of pairs is never less than the number of triples. If the number of image and map points are fixed at n and m, respectively, then the expected value of $P - T$ can be computed from the expected value of $k = (k_2, . . ., k_m)$. In the case $m = 5$, we need only compute the expected value of (k_2, k_3, k_4).

The analytical results on triples matching are not yet adequate to enable us to draw conclusions on the efficiency of the use of triples for transformation screening. The experimental results on triples matching, given in Section 4.11, indicate that triples are likely to be highly effective in reducing the number of transformations which must be examined.

4.11 Matching Triples Experiment

The theoretical analysis of the effectiveness of matching triples of points to determine a registration transformation has not yet reached a stage where the expected number of matches can be found. A simulation study was conducted to estimate the number of matching triples for fixed map size and varying image size.

The simulation consisted of twenty trials divided into four groups of five trials each. For each trial a map containing five points was generated using a uniform random number generator for the x and y coordinates of all points. A perturbed version of each point was then copied into the map where the perturbation consisted of adding random numbers generated by a mean zero, variance one normal process to the x and y coordinates. Random noise points were then added to the image. In each group of five trials, the number of image and map points were fixed. Only the locations of image and map points varied. The number of image points over all trials ranged from 10 to 25 in increments of five points.

Two types of matching were computed for each trial, the number of matching pairs and the number of matching triples. A matching pair ((a,b),(c,d)) is a pair of map points a and b and a pair of image points d and e such that the distance from a to b differed by no more than five from the distance from d to e. A matching triple is a set of three map points a, b, and c and three image points d, e, and f such that the following are matching pairs:

((a,b),(d,e))

((a,c),(d,f))

((b,c),(e,f))

Thus the triangles formed by the two triples are approximately congruent.

The results of the simulation are given in Table 4.7. For each image map pair, the ratio of the number of matching pairs to the number of matching triples was computed. For each group of

five trials in which the number of image points was constant, the five ratios were arranged. The average ratios are approximately a linear function of the number of points. In the range of parameter values studied, we may state that the following relationship holds approximately:

$$R = .1N + 1.19,$$

where R is the ratio of the number of matching pairs to the number of matching triples for N noise points in the image.

The above results have several applications. First, in the range of parameter values studied, the selection of matching triples reduces the number of possible transformations sufficiently to permit exhaustive enumeration. On an image with 20 noise points, 5 map points, and 100 image edges randomly selected, the registration with the clustering program would often evaluate 20 to 30 transformations explicitly while the triples screening procedure produced an average of 32 transformations to examine for the same type of image. Thus the triples screening provided only slightly less data reduction than the clustering. In view of the limitations on clustering discussed in Section 5, this suggests replacement of clustering by triples screening if this can be done rapidly enough.

A second useful piece of information gained from this study is the approximate numerical relationship between the number of noise points and the ratio between the number of matching triples and matching pairs. In trying to design a fast registration algorithm, we would like to balance the cost of successful but relatively expensive transformation screening such as triple matching versus cheaper but less effective screening such as pair matching. If we can estimate the computational cost of pair matching and triple matching as a function of the number of noise points, the approximate equation allows us to assign to each number of noise points the expected number of transformations and

the expected cost resulting from matching pairs and triples. This, in conjuntion with an estimate of the cost of evaluating a transformation, would allow us to determine the preferable screening method and estimate the average running time of the algorithm.

Table 4.7 Number of matching pairs (P) of points, number of matching triples (T) of points and ratio of P to T (P/T) for a five point map and a 5+N point image.

N	P	T	P/T
5	18	15	1.20
5	19	14	1.36
5	26	25	1.04
5	25	24	1.04
5	21	14	1.50
10	31	17	1.82
10	32	15	2.13
10	40	27	1.48
10	38	27	1.41
10	34	21	1.62
15	46	19	2.42
15	48	17	2.82
15	54	28	1.93
15	63	42	1.50
15	53	23	2.30
20	68	21	3.24
20	69	18	3.83
20	81	38	2.13
20	78	53	1.47
20	77	28	2.75

5. Clustering

The clustering algorithm in the LNK registration procedure, described briefly in Section 2, has provided reasonable results in many of the experiments described in this report and in earlier LNK studies [Stockman 1981]. Despite this we feel that in situations in which abstract vectors are used, the clustering can and should be avoided.

There are several reasons for avoiding clustering whenever possible. First, as seen in Section 4.6, the current clustering procedure may fail to detect a peak near the true transformation. This situation can arise in several ways. To avoid losing clusters due to clusters being spread over several bins without adequate concentration in any one bin, the clustering program uses overlapping bins. Unfortunately, more overlapping bins must be used to avoid this fragmentation. As the number of parameters to be estimated grows, as when we add scale and perspective changes, the number of overlapping bins grows rapidly, thus slowing down the clustering considerably. We do not yet have any information as to the frequency of occurrence of this phenomenon other than to observe that it has occurred.

The fragmentation may be avoided to some degree by searching for more peaks in the cluster space. At present, only three peaks are being used. More peaks may be considered, although this can increase the computational cost significantly when the clustering proceeds to deeper levels.

A second reason for avoiding the clustering is that the clustering cannot be easily adapted to an iterative registration scheme in which the clustering based on a set of points must be revised as new points are added (see Section 6). It is possible that iterative clustering may work, but we think that more effort in screening out infeasible transformations is preferable to the use of clustering for picking out likely transformations.

We think that analysis of the clustering algorithm cannot effectively be carried out until information has been gathered about the effects in real data of sampling and labelling as

outlined in Section 4. The theoretical analysis of clusterinng procedures has produced almost no useful results in the clustering literature, so if clustering is necessary, analysis will be primarily empirical.

6. Iterative Registration

The high cost of feature point extraction suggests consideration of registration procedures which perform feature extraction on a limited portion of the image and attempt registration based on the output. If the results are not satisfactory, additional portions of the image should be examined and the registration should be attempted once again. This process would continue until either an acceptable registration is achieved or the entire image has been examined for point features.

Three basic issues must be resolved before this iterative type of procedure can be made feasible. First, a selection procedure must be designed to determine which portions of the image are examined on successive iterations. Second, the procedure for determining which registration transformations to examine should be modified to allow the procedure to use results from previous steps of the iteration. Third, a termination criterion must be designed for acceptance or rejection of a registration transformation determined after an iteration of the procedure.

There are two main classes of procedures for selecting which portion of the image should be examined after an iteration. First, the image can be partitioned into square windows. An unexamined window could be selected at random after each iteration. In the second class of methods, the best registration transformation from the previous iteration can be used as an estimate of the true transformation. If an image is being registered to a map, this estimate can then be used to determine where other map feature points are likely to appear in the image. A program for performing this second method is currently under development, though in light of other results in this report, we feel the method of random window selection appears more fruitful, and attention should be turned in this direction.

The main reason for preferring random window selection is that the LNK procedure tends to produce transformations which are

either fairly accurate, say to within a few pixels and a few degrees, or transformations which are very far off (see Section 4.6). Thus the estimated transformation in the second method is likely to be sufficiently bad as to be worthless or sufficiently good that no further iterations are required. For fine tuning the registration or verification of the registration transformation, the registration transformation may be used to seek image features corresponding to map features, but this is a one-step process.

The modification of the registration procedure to use information from previous iterations to determine which transformations to examine can be handled in two ways. First, some sort of sequential clustering procedure could be defined. Such a procedure would cluster a set of transformations and then update the clustering when a new set of transformations were added. This type of approach appears risky since updating procedures would tend to look at characterizations of the points being clustered rather than the points themselves. These characterizations may include cluster centers and radii. Thus the updated clusters may be much less reliable than a direct clustering on all points.

A second procedure for using information from previous transformations is available if the number of feasible transformations is sufficiently small that exhaustive search is possible. Based upon the results of Section 4, we think that efforts should be concentrated in this direction. As new feature points are added on successive iterations, the updating of the list of transformations is a simple matter. Procedures for updating the match weight assigned to transformations are more tricky, but we think they can be rapid.

We have not yet investigated termination criteria for iterative registration. It was felt that more experience on real data using the inersection detection program should be obtained before reasonable criteria can be set.

7. Conclusions and Recommendations

This study was directed toward the analysis of the LNK registration procedure, both to understand its performance in its present form and to investigate modifications which may result in improved registration accuracy and speed. For purposes of analysis, we have focused on theoretical methods and simulation to gain knowledge of the relative merits of various facets of the algorithm. We now summarize the primary results of this work.

The curve intersection detection algorithm of Section 3.2 appears to be a feasible approach to the extraction of feature points, thus enhancing the straight line intersection detector currently available. This procedure should now be tested on aerial imagery to obtain information on characteristics of the feature points obtained.

The LNK procedure, viewed as an approximation to a correlation-based method for registration of point images, has remarkably high accuracy in locating the correct transformation. The investigations described in Sections 4.2, 4.3, and 4.4 provide both empirical and analytical evidence for this conclusion.

According to the study in Section 4.6, there are two main causes of registration error using the LNK procedure. First, the clustering routine may fail to detect the true transformation due to the cluster's containing this transformation being split over several bins. Second, the number of correctly matching vectors may be too small to give a high match weight to the correct transformation.

The second problem may be alleviated by more elaborate screening including using matching triples of points, labels of points, selection of map vectors based on good length properties, and other methods described in Sections 4.7 - 4.11. Each of these methods can significantly reduce the number of vectors to be handled by the current registration program, thus increasing the expected ratio of correctly matching vectors to all input vectors. In the range of parameter values considered, several of

these procedures provided enough reduction in the number of matching vectors to allow elimination of the clustering program, thereby eliminating one source of error.

Combinations of screening methods such as using labelled matching triples of points were not studied, but based on the present work, they appear likely to improve registration accuracy still further. The potential of such combinations is great, but the computational cost of examining triples must be examined carefully, for its worst-case performance can deteriorate rapidly as the number of map points increases.

The performance of the registration algorithm is dependent upon the way in which the map feature points are selected from the map. It is both feasible and desirable to select map vectors so no two are similar in length. Further conditions on map vector selection can be devised to increase the likelihood of the registration procedure providing an accurate registration transformation.

Further study of the LNK registration procedure should include both analytical and experimental research. The most pressing need is the acquisition of feature point images derived from aerial imagery using our intersection detection routines. Maps of the corresponding areas would be useful. These point images would allow extensive testing of the ideas developed in this report.

The present analysis has focussed on the estimation of rotation and translation parameters using the LNK procedure. The introduction of scale and perspective distortions leads to greater complexity which may require different methods of analysis. In particular, clustering may be essential in this more general situation. A study of the extent to which complexity is increased by scale and perspective distortions would be valuable.

New methods for reducing the number of matching vectors or

triangles should be studied in the case where scale and perspective distortions are present. A study should be made of properties of geometric figures which are invariant under these two types of distortions. This would aid us in designing procedures for recognizing geometric configurations which could not match under any scale or perspective change. This could remove much of the load from the clustering routine.

The random sampling of the image suggested in Section 6 should now be explored in detail. Criteria for accepting a tentative registration should be designed and tested. Updating algorithms for determining match weights and feasible transformations should also be designed.

Simulation experiments with 3-dimensional matching should be conducted. Attention should be given to the potential obscuring of feature points due to viewing angle. The creation of artificial terrain models would aid us in performing experimentation along these lines.

References

M.A. Fischler and R.C. Bolles, "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography," CACM, Vol. 24, June 1981, pp. 381-395.

N. Johnson and S. Kotz, Discrete Distribution, Houghton-Mifflin Co., Boston, 1969.

B.A. Lambird, D. Lavine, G.C. Stockman, K.C. Hayes, and L.N. Kanal, Study of Digital Matching of Dissimilar Images, Report No. ETL-0248. Final report by L.N.K. Corp. under Contract DAAK 70-79-C-0234 to U.S. Army Engineering Topographic Laboratories, Fort Belvoir, Virginia, 1980.

H. Mostafavi, "Optimal Window Functions for Image Correlation in the Presence of Geometric Distortions," IEEE Transactions on Acoustics, Speech and Signal Processing, Vol. ASSP-27, No. 2, April 1979, pp 163-169.

G.C. Stockman and S.H. Kapstein, "The Use of Models in Image Analysis," Report No. AMRL-TR-78-117. Final Report by L.N.K. Corp. under Contract F33615-76-C-0521 to Aerospace Medical Research Laboratory, Wright-Patterson A.F.B., Ohio, January 1979.

G.C. Stockman, B.A. Lambird, D. Lavine, L.N. Kanal, Knowledge-Based Image Analisis, Report No. ETL 0258. Final report by L.N.K. Corp. under contract DAAK-70-77-C-0110 to U.S. Army Engineering Topograhic Laboratories, Fort Belvoir, Virginia, 1981.

DAT
FILM